

Genealogies in simple models of evolution

Éric Brunet and Bernard Derrida

Laboratoire de Physique Statistique, École Normale Supérieure,
UPMC, Université Paris Diderot, CNRS,
24, rue Lhomond, 75231 Paris Cedex 05, France

E-mail: Eric.Brunet@lps.ens.fr

Abstract. We review the statistical properties of the genealogies of a few models of evolution. In the asexual case, selection leads to coalescence times which grow logarithmically with the size of the population in contrast with the linear growth of the neutral case. Moreover for a whole class of models, the statistics of the genealogies are those of the Bolthausen Sznitman coalescent rather than the Kingman coalescent in the neutral case. For sexual reproduction, the time to reach the first common ancestors to the whole population and the time for all individuals to have all their ancestors in common are also logarithmic in the neutral case, as predicted by Chang [1]. We discuss how these times are modified in a simple way of introducing selection.

PACS numbers: 02.50.-r, 05.40.-a, 87.10.-Mn, 89.75.Hc

1. Introduction

The genealogy of a population describes the relationships between all the ancestors of this population. Simple questions one may ask about the genealogy of a population are:

How far one has to go into the past to find the most recent ancestor of two individuals?
of m individuals? of the whole population?

How do these times depend on the sample of m individuals chosen at random in the population?

How do they depend on the size of the population?

How do they fluctuate when the population evolves over many generations?

How are they affected by the forces (like selection) acting on evolution?

In the case of *an asexual evolution*, the ancestry of a population is a tree, the root of which is the most recent common ancestor of the whole population. In the neutral case (i.e. when all individuals have on average the same number of surviving offspring at the next generation), for a well mixed population, the height of the tree is proportional to the size of the population (see figure 1) and its statistics are described by Kingman's coalescent [2, 3, 4, 5] (see section 2).

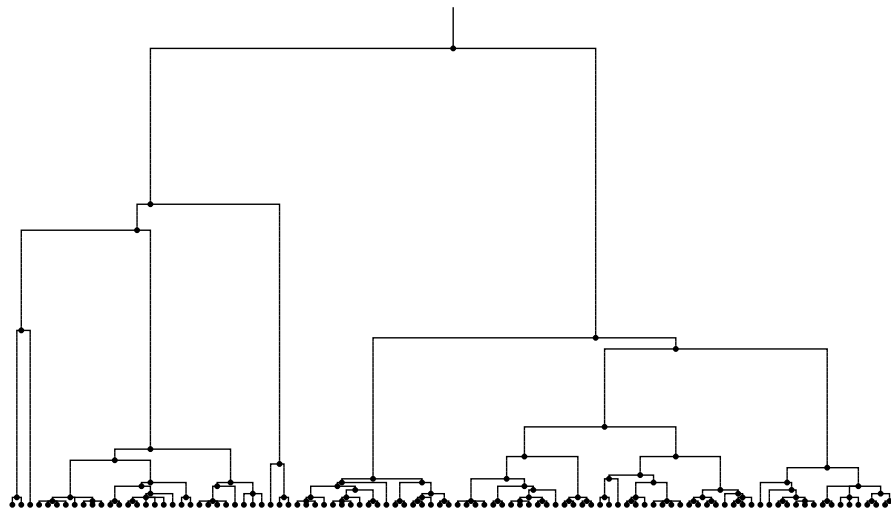


Figure 1. A typical genealogical tree in the neutral case, obtained by simulating the Wright-Fisher model for a population of $N = 100$ individuals. For the particular realization shown on the figure, the number of generations to reach the most recent common ancestor is 125, which is, as expected, of order N . Already for $N = 100$, all the visible nodes are coalescences of pairs of branches and one cannot see any multiple coalescences except at the very bottom of the figure where the number of branches is still of order N .

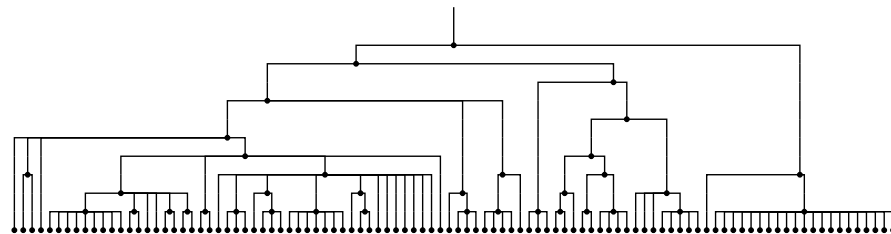


Figure 2. A typical genealogical tree in presence of selection for a population of size $N = 100$. The number of generations (10 for the realization of the exponential model shown on the figure) to reach the most recent common ancestor is much shorter than in the neutral case. In contrast to the neutral case, one can observe multiple coalescences even rather high in the tree.

In a series of recent papers, together with A.H. Mueller and S. Munier [6, 7], we considered a family of models of evolution with selection. For these models, in contrast to the neutral case, the height of the tree grows logarithmically with the size of the population and, its shape (see figure 2) is given asymptotically by the Bolthausen-Sznitman coalescent [8] rather than by Kingman's coalescent.

In the case of *sexual evolution*, each individual has two parents, four grand parents, and so on. Each individual is therefore the root of what looks like a tree for the first generations in the past. Going further into the past, however, the branches of this tree start to merge (see figure 3) and the number of distinct ancestors do not grow exponentially anymore [1, 9]. The number of ancestors then saturates at a value which is a fraction of the whole population living in this remote past (the rest of this past

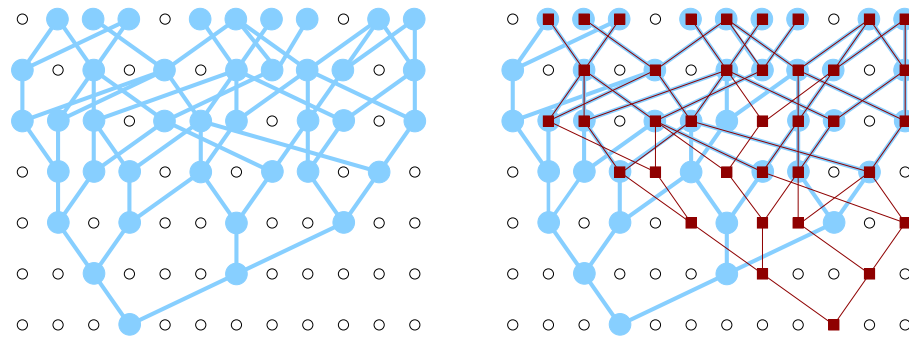


Figure 3. The “tree” of ancestors in the case of sexual reproduction. After a few generations (of order the log of the size of the population), the branches start to merge and the total number of ancestors saturates (left figure). Furthermore, comparing the ancestries of two individuals (right figure), all their ancestors become identical.

population consists of all those individuals who had no offspring or whose lineage became extinct). Comparing the ancestry of two individuals of the same generation, one sees two growing binary trees in the recent past which then start to intermix in a more remote past until they become identical.

Here we try to review a few properties of the genealogies of some simple models of evolving populations, both in the neutral case and in the case of selection. In section 2, we recall a few properties of the Kingman coalescent and of the Bolthausen-Sznitman coalescent, as special cases of coalescent processes and why neutral evolution in the asexual case leads to the Kingman coalescent, in the limit of a large population. In section 3, we describe the properties of the exponential model, an exactly soluble of asexual evolution in presence of selection, and show that its trees follow the statistics of the Bolthausen-Sznitman coalescent. In section 4, we show that, like the exponential model, more generic models of evolution with selection converge to the Bolthausen-Sznitman coalescent. Section 5 is devoted to the case of sexual evolution both in the neutral case and in presence of selection.

2. How to quantify genealogies in the case of asexual reproduction

To study the statistical properties of the trees generated by sexual reproduction, one can adopt several points of view:

- (i) One can try to see how the population can be partitioned into τ -families, with the rule that two individuals belong to the same τ -family if and only if their most recent common ancestor is at a distance less than τ generations in the past. These τ -families can themselves be divided into subfamilies by choosing a shorter number of generations τ' , and so on.
- (ii) One can alternatively study the genealogical tree as a dynamical process and try to determine the rates q_k at which k branches merge into one when one moves from the bottom to the top of the tree (*i.e.* when one looks at the tree backwards in

time). These rates might be correlated in time or depend on the characteristics of the individuals along the branches which merge.

- (iii) In our previous works [6, 7, 10] we chose to compute for these random trees the coalescence times T_p defined as follows: T_p is the age of the most recent common ancestor of p individuals chosen at random in the population. In general the times T_p depend both on the sample of p individuals chosen and on the generation. One advantage of characterizing the random trees by the times T_p is that they are relatively easy to average (over the samples of p individuals and over the generations) in simulations.

For some models, such as coalescence processes discussed below, one can find explicit mathematical formulae which relate these various properties [10].

2.1. Coalescence models

Coalescent processes give a simple procedure to generate a whole class of random trees [11, 12, 13, 14]. In a continuous time version, a coalescent process is a dynamical stochastic process, where any set of k individuals have a probability $q_k dt$ of coalescing into one individual during the infinitesimal time interval dt , implying that, if the number of branches is b at time t , the probability $r_b(b') dt$ of having b' branches at time $t + dt$ (with $b' < b$) is given by [10]

$$r_b(b') = \sum_{n=0}^{b'-1} \frac{b!}{(b-b'+1)! (b'-1-n)! n!} (-1)^n q_{n+b-b'+1}. \quad (1)$$

From the rates q_k , one can easily calculate the times $\langle T_p \rangle$ (averaged over all the realizations of the coalescence process) by analyzing what happens in a steady state situation during an infinitesimal time interval dt

$$\langle T_p \rangle = dt + \langle T_p \rangle \left(1 - dt \sum_{b' < p} r_p(b') \right) + dt \sum_{2 \leq b' < p} r_p(b') \langle T_{b'} \rangle + O(dt^2). \quad (2)$$

One then gets

$$\begin{aligned} \langle T_2 \rangle &= \frac{1}{q_2}, & \frac{\langle T_3 \rangle}{\langle T_2 \rangle} &= \frac{4q_2 - 3q_3}{3q_2 - 2q_3}, \\ \frac{\langle T_4 \rangle}{\langle T_2 \rangle} &= \frac{27q_2^2 - 56q_2q_3 + 28q_3^2 + 12q_2q_4 - 10q_3q_4}{(3q_2 - 2q_3)(6q_2 - 8q_3 + 3q_4)}. \end{aligned} \quad (3)$$

So q_2 determines the time scale of the times T_p and all the ratios $\langle T_p \rangle / \langle T_2 \rangle$ are given by ratios of the coalescing rates q_k .

Remark: an alternative way of thinking of the coalescence processes [12, 14, 15] defined above is to say that during every infinitesimal time interval dt , there is a probability $\rho(f)dfdt$ that a fraction f of all the branches coalesces into a single branch, all the other branches remaining unchanged. Then for k given individuals, the probability that their k branches merge during dt is given by

$$q_k dt = \int_0^1 f^k \rho(f) df dt, \quad (4)$$

while the probability $r_b(b')dt$ that b distinct branches becomes b' branches during the time interval dt is

$$r_b(b')dt = \frac{b!}{(b'-1)!(b-b'+1)!} \int_0^1 f^{b-b'+1}(1-f)^{b'-1} \rho(f) df dt \quad (5)$$

and one recovers (1).

Two special cases will be of interest in what follows:

(i) *The Kingman coalescent*

In the Kingman coalescent, only $q_2 \neq 0$ and all the $q_k = 0$ for $k \geq 3$. Then (3) becomes

$$\langle T_2 \rangle = \frac{1}{q_2}, \quad \frac{\langle T_3 \rangle}{\langle T_2 \rangle} = \frac{4}{3}, \quad \frac{\langle T_4 \rangle}{\langle T_2 \rangle} = \frac{3}{2}. \quad (6)$$

In fact all the correlation functions between these times T_p can be computed [16, 5, 12, 17]. In particular one gets $\langle T_p \rangle / \langle T_2 \rangle = 2 - 2/p$.

(ii) *The Bolthausen-Sznitman coalescent*

The Bolthausen-Sznitman coalescent [8] was introduced in the context of the mean field theory of spin glasses [18, 19] to represent the tree structure of the pure states [20, 21] predicted by the replica scheme invented by Parisi [22, 23, 24]. In terms of the rates q_k , the Bolthausen-Sznitman coalescent can be defined by

$$q_k = \frac{q_2}{k-1} \quad (7)$$

leading to

$$\langle T_2 \rangle = \frac{1}{q_2}, \quad \frac{\langle T_3 \rangle}{\langle T_2 \rangle} = \frac{5}{4}, \quad \frac{\langle T_4 \rangle}{\langle T_2 \rangle} = \frac{25}{18}. \quad (8)$$

A natural question, then, is to know whether a given model of evolution (with or without selection) gives rise to random trees which can be described by a particular coalescent process [11, 14, 7, 15, 10].

2.2. The Wright-Fisher model

The Wright-Fisher model [11, 25, 14, 26] is one of the simplest and most studied models of evolution in the neutral case. In its simplest version, it describes a population of fixed size N . At every generation g , the parent of each individual is chosen uniformly among the N individuals living at the previous generation. It is easy to see that the probability q_k that k individuals have the same parent at the previous generation is $q_k = N^{1-k}$, so that $q_k \ll q_2$ for $k > 2$ and large N . Moreover the probability of seeing more than one coalescence event among k individuals (with k of order 1) at a given generation becomes also much smaller than q_2 for large N . Therefore in the limit of a large population, the time $\langle T_2 \rangle = q_2^{-1}$ scales like N

$$\langle T_2 \rangle \sim N \quad (9)$$

the ratios $q_k/q_2 \rightarrow 0$ for $k \geq 3$, and the statistics of the trees are given by the Kingman coalescent.

The Kingman coalescent is particularly central in the theory of neutral evolution because it is universal [25, 16, 14, 5, 11]: one can change the rules in the definition of the Wright-Fisher model in many ways and in the limit of a large population, one always recovers the Kingman coalescent. For example one may choose the parent i in the previous generation with a non uniform probability p_i and as long as these p_i decay fast enough with the size N of the population, one has in the large N limit $q_k/q_2 = (\sum_i p_i^k)/(\sum_i p_i^2) \rightarrow 0$ for $k \geq 3$ and one recovers the Kingman coalescent [27].

Remark 1: Starting from the random tree structure of the Kingman coalescent, and assuming mutations arising at constant rate along the branches of these trees, one can predict the statistical properties of the genetic diversity [28, 29, 30, 31, 32, 33, 34].

Remark 2: Even in the large N limit, the shape of the tree and the times T_p fluctuate when one follows the same population over many generations. The correlations between these times at different generations can be computed [35, 17]. These fluctuating shapes of the trees are very reminiscent of the tree structure of pure states predicted by the mean field theory of spin glasses [20, 21, 32, 36].

Remark 3: One can define a finite dimensional version of the Wright-Fisher model by considering that the individuals of the population are at the nodes of a lattice of L^d sites with periodic boundary conditions in dimension d and that each individual has its parent chosen at random among all the sites at a distance less than l with $l \ll L$. [37, 38]. In dimension $d \geq 2$, the genealogies of such populations have their statistics still given by the Kingman coalescent (the problem can be formulated as a reaction diffusion problem whose upper critical dimension is 2). On the other hand, in dimension 1, the statistical properties of the trees are modified (they can be understood in terms of coalescing random walks) and (6) is then replaced by [39]

$$\langle T_2 \rangle \sim N^2, \quad \frac{\langle T_3 \rangle}{\langle T_2 \rangle} = \frac{7}{5}, \quad \frac{\langle T_4 \rangle}{\langle T_2 \rangle} = \frac{8}{5}. \quad (10)$$

Note that in this case the spatial aspect is crucial, and there is no choice of the q_k for which the one dimensional problem could be reduced to a coalescent process as defined in section (2.1): for all the coalescent processes, the age T_2 of the most recent common ancestor of two individuals has an exponential distribution implying that $\langle T_2^2 \rangle / \langle T_2 \rangle^2 = 2$ whereas [39], for the finite dimensional model in $d = 1$, this ratio takes the value $12/5$.

3. The exponential model: an exactly soluble model of evolution with selection

3.1. Definition

The exponential model [6, 7, 10] is a simple generalization of the Wright-Fisher model, which includes the effect of selection. As in the Wright-Fisher model the size N of the population remains the same at every generation. All the individuals at a given generation are however not equivalent: each individual i at generation g carries a value $x_i(g)$ which represents a trait (or a fitness in the sense used by Bak and Sneppen [40]). There are then two steps to go from generation g to generation $g + 1$:

The reproduction step: each individual has its offspring generated by a Poisson process of density $e^{-(x-x_i(g))}$. This simply means that, with probability $e^{-(y-x_i(g))}dy$, there is an offspring of $x_i(g)$ in the infinitesimal interval $(y, y + dy)$. Repeating the procedure for all individuals i at generation g , one gets that way an infinite number of offspring along the real axis, with many offspring for large negative y and no offspring for sufficiently large positive y (note that at the right of any point y on the real axis there are always a finite number of offspring).

The selection step: at generation $g + 1$, one keeps only the N rightmost points among all these offspring.

The simplicity of the exponential model comes from the following identity

$$\sum_{1 \leq i \leq N} e^{-(x-x_i(g))} = e^{-(x-X_g)} \quad \text{with} \quad X_g = \ln \left[\sum_{1 \leq i \leq N} e^{x_i(g)} \right]. \quad (11)$$

This means that the offspring of all the individuals at generation g can be generated by a single Poisson process centered at position X_g . So X_g is the only information about generation g needed to generate the next generation. A simple consequence is that the successive shifts $X_{g+1} - X_g$ of the position of the population are i.i.d. variables. Another advantage of the exponential model is that the population at generation g consists of the N rightmost points of a single Poisson process (11) centered at X_g . One can then show [10] that the population at generation $g + 1$ can be generated in the following way:

$$x_i(g + 1) = X_g + Y_{g+1} + y_i(g + 1), \quad (12)$$

where Y_{g+1} and the N variables $y_i(g + 1)$ are independent random variables distributed according to

$$\begin{aligned} \text{Prob}(Y < Y_{g+1} < Y + dY) &= \frac{1}{N!} \exp \left[-(N + 1)Y - e^{-Y} \right] dY, \\ \text{Prob}(y < y_i(g + 1) < y + dy) &= \begin{cases} e^{-y} dy & \text{for } y > 0, \\ 0 & \text{for } y < 0. \end{cases} \end{aligned} \quad (13)$$

The positions $\{x_i(g)\}$ form a cloud of points which remain grouped as g increases. One could locate this cloud of points by its center of mass $X_g^{\text{C.M.}}$. As the points remain

grouped, one can as well use the position X_g defined in (11) which has the advantage that the differences

$$X_{g+1} - X_g = Y_{g+1} + \ln \left[\sum_i e^{y_i(g+1)} \right] \quad (14)$$

are uncorrelated random variables. It is easy to see that the difference

$$X_g - X_g^{\text{C.M.}} = \ln \left[\sum_i e^{y_i(g)} \right] - \frac{1}{N} \sum_i y_i(g) \quad (15)$$

has a distribution independent of g , so that all the cumulants of $X_g^{\text{C.M.}}$ and of X_g have the same linear growth with g .

For example the speed of adaptation v_N (which is simply the velocity of the cloud of points along the real axis) is given by

$$v_N = \lim_{g \rightarrow \infty} \frac{\langle X_g \rangle}{g} = \lim_{g \rightarrow \infty} \frac{\langle X_g^{\text{C.M.}} \rangle}{g} = \langle X_g - X_{g-1} \rangle = \langle Y_g \rangle + \left\langle \ln \left[\sum_i e^{y_i(g)} \right] \right\rangle. \quad (16)$$

Similarly the generating function $G_N(\beta)$ of all the cumulants of the positions X_g or $X_g^{\text{C.M.}}$ can be computed by

$$\begin{aligned} G_N(\beta) &= \lim_{g \rightarrow \infty} \frac{\ln \langle e^{-\beta X_g} \rangle}{g} = \lim_{g \rightarrow \infty} \frac{\ln \langle e^{-\beta X_g^{\text{C.M.}}} \rangle}{g}, \\ &= \ln \langle e^{-\beta(X_g - X_{g-1})} \rangle = \ln \langle e^{-\beta Y_g} \rangle + \ln \left\langle \left[\sum_i e^{y_i(g)} \right]^{-\beta} \right\rangle. \end{aligned} \quad (17)$$

The average over Y_g and the $y_i(g)$'s (distributed according to (13)) can be performed leading to the following large N behavior [6, 7]

$$G_N(\beta) \simeq -\beta \ln \ln N - \frac{1}{\ln N} \left(\beta \ln \ln N + \beta \Gamma'(1) + \beta - \beta \frac{\Gamma'(1 + \beta)}{\Gamma(1 + \beta)} \right) + \dots \quad (18)$$

This gives for large N

$$v_N = \lim_{g \rightarrow \infty} \frac{\langle X_g \rangle}{g} = \ln \ln N + \frac{\ln \ln N + 1}{\ln N} + \dots \quad (19)$$

$$\lim_{g \rightarrow \infty} \frac{\langle X_g^k \rangle_c}{g} = \frac{k!}{\ln N} \left[\sum_{i \geq 1} \frac{1}{i^k} \right] + \dots \quad (20)$$

This logarithmic growth of the velocity with the size of the population seems to be common to many models of evolution with selection on a smooth landscape [41, 42, 43] where each individual has a number of offspring proportional to the exponential of its fitness.

3.2. The genealogies in the exponential model

In the exponential model, given that at generation $g+1$ there is an individual at position x , the probability $p_i(x, g+1)$ that this individual is the offspring of the i -th individual at generation g is given by

$$p_i(x, g+1) = \frac{e^{-(x-x_i(g))}}{\sum_j e^{-(x-x_j(g))}} = \frac{e^{x_i(g)}}{\sum_j e^{x_j(g)}} = \frac{e^{y_i(g)}}{\sum_j e^{y_j(g)}}. \quad (21)$$

The simplicity of the exponential model comes from the fact that these p_i do not depend on x and that they can be expressed in terms of the $y_i(g)$ only, implying that the p_i at different generations are uncorrelated.

The probability q_k that k individuals have a single common ancestor at the previous generation is then given by

$$q_k = \left\langle \sum_i p_i^k \right\rangle = \left\langle \frac{\sum_i e^{ky_i(g)}}{\left[\sum_j e^{y_j(g)}\right]^k} \right\rangle, \quad (22)$$

where the average is over the N random variables $y_i(g)$ distributed according to (13). For large N these averages can be computed [10]

$$q_k \simeq \frac{1}{(k-1) \ln N}. \quad (23)$$

Moreover, at each generation, the probability of more than one coalescence event among a fixed number of individuals becomes negligible in the large N limit [10]. Therefore the exponential model converges to the Bolthausen-Sznitman coalescent (7) with coalescence times given by (8) and

$$\langle T_2 \rangle \sim \ln N. \quad (24)$$

3.3. Conditionning on the speed

The shapes and the heights of the trees fluctuate with g . In [10] we tried to determine how these fluctuations are correlated to the displacement $X_g - X_0$. To do so we assigned to each evolution event a weight $e^{-\beta X_g}$. This means that, if $\overline{T_k(g)}$ is the age of the most recent ancestor averaged over all the choices of k individuals at generation g (for one realization of the process), we tried to compute weighted averages such as

$$\langle T_k \rangle_\beta = \lim_{g \rightarrow \infty} \frac{1}{g} \sum_{g'=1}^g \frac{\langle e^{-\beta X_{g'}} \overline{T_k(g')} \rangle}{\langle e^{-\beta X_{g'}} \rangle}. \quad (25)$$

One can then show [10] that, with these weights, the probability that k individuals at generation $g+1$ have their most recent common ancestor at the previous generation is

$$q_k = \frac{\left\langle \sum_i e^{ky_i(g)} \left[\sum_j e^{y_j(g)}\right]^{-\beta-k} \right\rangle}{\left\langle \left[\sum_j e^{y_j(g)}\right]^{-\beta} \right\rangle}. \quad (26)$$

As for $\beta = 0$, for large N , the probability of observing more than one coalescence event at each generation becomes negligible and (26) gives to leading order [10]

$$q_k = \frac{1}{\ln N} \frac{(k-2)! \Gamma(\beta+1)}{\Gamma(\beta+k)}. \quad (27)$$

Therefore for these biased events, the trees have the same statistics as a coalescent whose rates are given by (27). This allows one to determine through (3) the coalescence times $\langle T_p \rangle$ (see [10] for explicit expressions).

Varying β in (27) we see that, conditioning on the displacement X_g , the statistics of the trees interpolate between the Bolthausen-Sznitman coalescent for $\beta = 0$ (no bias) and the Kingman coalescent for $\beta \rightarrow \infty$ (bias toward slow velocities).

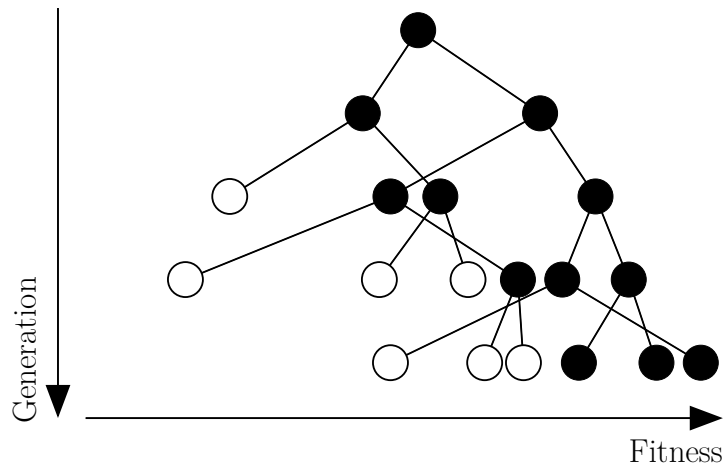


Figure 4. The perfect selection case: at each generation one selects the N (here $N = 3$) rightmost offspring produced by the previous generation.

4. More generic models of evolution with selection

In [6, 7, 10], we considered a whole family of models of evolution with selection which generalize the exponential model. These more generic models are defined as follows: as in the exponential model, the population has a fixed size N and each individual i at generation g is characterized by a real number (a trait) $x_i(g)$. At the next generation, the trait $x_i(g+1)$ of an offspring i is inherited from its parent $P(i, g+1)$ up to a random number $z_i(g+1)$ drawn from a fixed distribution $\rho(z)$ which represents the effect of the mutations on the trait x_i

$$x_i(g+1) = x_{P(i,g+1)}(g) + z_i(g+1). \quad (28)$$

Then comes the selection step where only N survivors among all the offspring produced by generation $g+1$ are chosen according to their trait $x_i(g+1)$.

The model [6, 7] depends on the number of offspring of each individual, on the distribution $\rho(z)$, and on the way the N survivors at generation $g+1$ are chosen among all the offspring of generation g .

Here we present the result of numerical simulations of three versions of the model, where we always took for $\rho(z)$ a flat distribution:

- *The perfect selection case:* in this version (see figure 4), each individual at generation g has two offspring, but out of the $2N$ resulting individuals we only keep at generation $g+1$ the N rightmost ones *i.e.* those who have the highest $x_i(g+1)$.
- *The fuzzy selection case:* here again, each individual has two offspring at the next generation, but out of the $2N$ resulting individuals, the N survivors at generation $g+1$ are chosen uniformly among the $3N/2$ rightmost ones.
- *The two parent selection case:* in this version, for each individual at generation $g+1$, instead of choosing its parent uniformly in the population at generation g

as in the Wright-Fisher model, we choose *two* potential parents uniformly in the population and keep as the actual parent the best one among these two [44]. (It is equivalent to state that the i -th best individual has on average $2(N - i)/(N - 1)$ offspring and then draw the new generation with a multinomial distribution.)

We measured the average coalescence times $\langle T_2 \rangle$, $\langle T_3 \rangle$ and $\langle T_4 \rangle$. The time $\langle T_2 \rangle$ seems to increase like $\ln^3 N$ (see figure 5) and the ratios of these times seem to converge to the values predicted for the Bolthausen-Snitzman coalescent in the large N limit, (see figure 6).

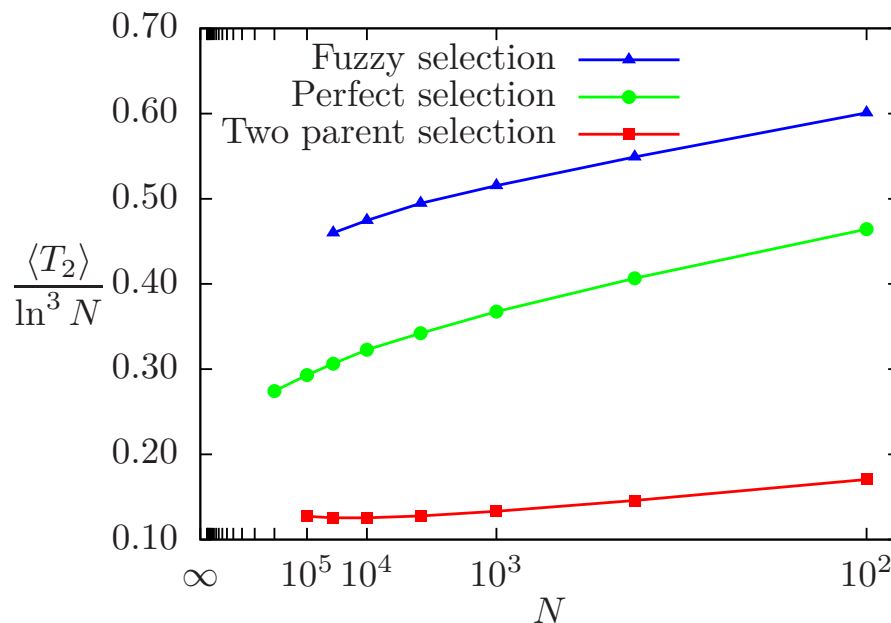


Figure 5. $\langle T_2 \rangle / \ln^3 N$ as a function of N for the three models. The scale on the horizontal axis is $1/\ln^2 N$.

The above numerical simulations and additional ones (some on larger scales) [6, 7, 10, 45] indicate that for this whole family of models the statistics of the genealogical trees converge, in the large N limit, to those of the Bolthausen-Sznitzman coalescent (8) as for the the exponential model, the only difference being that the coalescence times $\langle T_2 \rangle$ grow like

$$\langle T_2 \rangle \sim \ln^3 N. \quad (29)$$

instead of $\ln N$ for the exponential model (24).

At each generation, the N values $x_i(g)$ form a cloud of points which moves along the real axis as g increases. This motion is stochastic and we argued in [6, 7] that its evolution can be related, for large N , to that of a noisy Fisher-KPP equation in the weak noise limit. In our joint works with A.H. Mueller and S. Munier [6, 7] we used a phenomenological theory [45] developed for such travelling wave equations to explain the convergence to the Bolthausen-Snitzman coalescent and the $\ln^3 N$ timescale. From this phenomenological theory the following picture emerges: the evolution of the

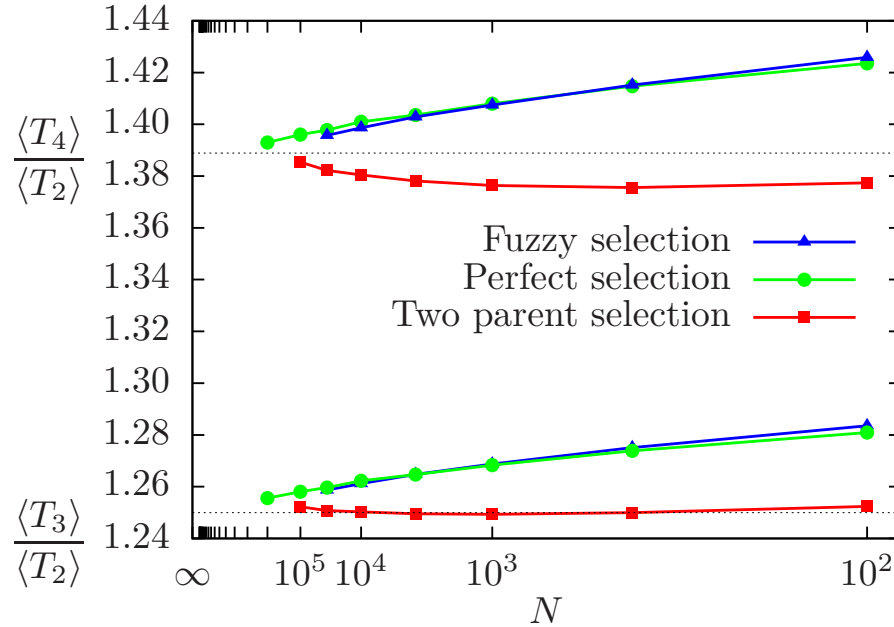


Figure 6. $\langle T_4 \rangle / \langle T_2 \rangle$ (up) and $\langle T_3 \rangle / \langle T_2 \rangle$ (down) as a function of N for the three models. The dotted lines are the values for the Bolthausen-Snitzman coalescent, respectively $25/18$ and $5/4$. The scale on the horizontal axis is $1/\ln^2 N$.

cloud of points (the $x_i(g)$) is most of the time deterministic. At time intervals of order $\ln^3 N$, rare fluctuations occur where the best individual gets an exceptionally good trait $x_i(g)$. These fluctuations relax on a time scale of order $\ln^2 N$ generations. During this relaxation time, the individuals with the exceptionally good trait have their long time lineage less affected by selection than the rest of the population and the net effect is that their lineage replaces a positive fraction f of the whole population as in (4); the value of f depends on the size of the fluctuation. An analysis of the distribution of these rare fluctuations leads to a distribution of f , in (4), which corresponds to the Bolthausen-Snitzman coalescent. So the picture resembles that of the model considered by Durrett and Schweinsberg [14] where the effect of selection ends up giving rise to a coalescent with multiple collisions.

It might look paradoxical to get the Bolthausen-Snitzman coalescent, with multiple branches coalescing at the same point, in situations like the perfect selection case, where each individual has only two offspring. This is simply because coalescences of pairs at nearby generations look like multiple coalescences on a time scale of order $\ln^3 N$. Figure 7 illustrates this mechanism. We simulated the model with perfect selection for a population of size $N = 10^4$ and, at an arbitrary time, we plotted the coalescing tree of the ancestries of 100 individuals chosen at random. While a close look makes it apparent that there are only coalescences of pairs, one can see several regions where coalescences of pairs occur in quick succession, thus resulting into the quasi-instantaneous coalescence of more than two lines [14]. Note also that during the $\ln^3 N$ generations (in real time) needed for a fluctuation to occur, the ancestral lines of the individuals diffuse within the

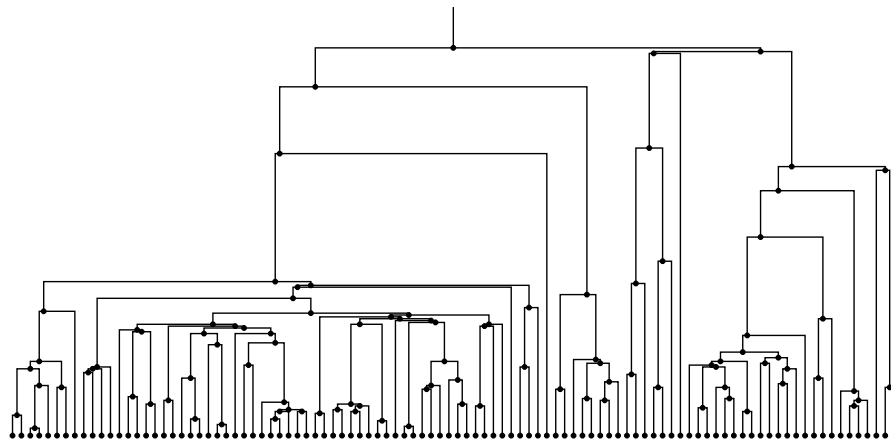


Figure 7. A realization of the genealogical tree of 100 individuals chosen at random among a population of 10^4 in the model with “perfect selection”.

whole population. As the size of this cloud is of order $\ln N$ along the real axis, it takes $\ln^2 N$ time steps to explore the whole population and by the time the next fluctuation occurs, these lines have swept many times over the whole population. This explains the perhaps counter-intuitive observation that the average time of coalescence between several individuals is independent of the positions of these individuals.

5. Sexual reproduction

In the case of sexual reproduction, each individual has two parents in the previous generation. When the population is large, its number n_g of distinct ancestors, g generations in the past, starts to grow like 2^g . For any finite population, however, this number n_g of distinct ancestors has to saturate and the branches of the genealogical tree have to merge if one goes far enough into the past (see figure 3). Looking at a given model of evolution, with sexual reproduction, for a population of fixed size, one may wonder :

- (i) How does n_g depend on g ?
- (ii) When does the saturation occur?
- (iii) How far one has to go into the past to find *some* common ancestors to a group of p individuals?
- (iv) When does a group of p individuals have *all* their ancestors in common?

5.1. The neutral case

In the neutral case, all these questions have simple answers. For a simple model discussed below, for example, Chang [1, 46] has shown that the number generations G_N for all individuals to have at least one common ancestor is (for large N)

$$G_N \sim \frac{\ln N}{\ln 2} \sim 1.44 \ln N, \quad (30)$$

while the number generations \widetilde{G}_N needed for the whole population to have *all* their ancestors in common is

$$\widetilde{G}_N \sim \left(\frac{1}{\ln 2} - \frac{1}{\ln(2(1-x^*))} \right) \ln N \sim 2.55 \ln N, \quad (31)$$

where $x^* \simeq 0.8$ is the non zero solution of $x^* = 1 - \exp(-2x^*)$.

The simplest model of neutral evolution to describe the genealogies in the case of sexual reproduction is to take a population of fixed size N and to assign to each individual two parents chosen at random in the previous generation (for simplicity we won't make any distinction between males and females; moreover, the two parents of each individual being chosen independently in the previous generation, the model allows these two parents to coincide with probability $1/N$. All these simplifications are in fact unimportant as they do not affect the large N behavior (30,31)). In such a model, the number n_g of distinct ancestors of a given individual, g generations ago in the past, is a Markov process. The distribution of n_{g+1} , given n_g , can be written but it leads to a rather complicated formula which is in fact not that useful to understand the large N behavior of the model. Instead one can calculate the first moments of n_{g+1} by writing that

$$n_{g+1} = N - \sum_{i=1}^N y_i, \quad (32)$$

where $y_i = 0$ if individual i at generation $g+1$ (in the past) is the ancestor of at least one of the n_g ancestors at generation g and $y_i = 1$ otherwise (i.e. $y_i = 1$ if i has no offspring among the n_g ancestors at generation g). For i, j, k distinct,

$$\langle y_i \rangle = \left(1 - \frac{1}{N}\right)^{2n_g}, \quad \langle y_i y_j \rangle = \left(1 - \frac{2}{N}\right)^{2n_g}, \quad \langle y_i y_j y_k \rangle = \left(1 - \frac{3}{N}\right)^{2n_g}, \quad (33)$$

etc. It is then easy to calculate the first moments of n_{g+1}

$$\langle n_{g+1} \rangle = N \left[1 - \left(1 - \frac{1}{N}\right)^{2n_g} \right], \quad (34)$$

$$\langle n_{g+1}^2 \rangle - \langle n_{g+1} \rangle^2 = N \left(1 - \frac{1}{N}\right)^{2n_g} + N(N-1) \left(1 - \frac{2}{N}\right)^{2n_g} - N^2 \left(1 - \frac{1}{N}\right)^{4n_g}. \quad (35)$$

For large N , one can see that

$$\langle n_{g+1} \rangle \simeq N \left(1 - \exp \left[-\frac{2n_g}{N} \right] \right), \quad (36)$$

and that

$$\langle n_{g+1}^2 \rangle - \langle n_{g+1} \rangle^2 \simeq \begin{cases} N \exp \left[-\frac{2n_g}{N} \right] - (N + 2n_g) \exp \left[-\frac{4n_g}{N} \right] & \text{for } n_g \sim N, \\ \frac{2n_g^2 - n_g}{N} & \text{for } n_g \ll N. \end{cases} \quad (37)$$

These expressions show that, for large N , the fluctuations of n_{g+1} are small compared to the average $\langle n_{g+1} \rangle$ and the evolution of n_g is very well approximated by a deterministic evolution

$$n_{g+1} = NF \left(\frac{n_g}{N} \right) \quad \text{and} \quad n_0 = 1, \quad \text{with} \quad F(x) = 1 - e^{-2x}. \quad (38)$$

As long as $n_g/N \ll 1$, the function $F(x)$ can be approximated by a linear function $F(x) \simeq 2x$, and one has $n_g \simeq 2^g$ for $g \ll \ln N$. On the other hand, in the long time limit, $n_g \rightarrow Nx^*$ where $x^* \simeq 0.8$ is the attractive fixed point of the map $x \rightarrow F(x)$. The meaning of x^* is simply the fraction of the population whose lineage does not become extinct after many generations, while $1 - x^*$ is the fraction of the population whose lineage becomes extinct.

The map $x \rightarrow F$ allows one to determine several other properties of the genealogies. For example if one tries to compare the genealogies of k individuals and calls $n_g(k)$ the total number of distinct ancestors of at least one of these k individuals at generation g in the past, it is clear that up to a change of initial condition $n_g(k)$ evolves as n_g in (38):

$$n_{g+1}(k) = NF\left(\frac{n_g(k)}{N}\right) \quad \text{with} \quad n_0(k) = k. \quad (39)$$

Again, as long as $g \ll \ln N$, one finds $n_g(k) = k 2^g$ and the ancestors of the k individuals are all distinct. On the other hand, all the $n_g(k)$ converge to the same value Nx^* meaning that all ancestors become common to the whole population.

The number of ancestors $m_g(2)$ common to 2 individuals at generation g in the past can be written as

$$m_g(2) = 2n_g(1) - n_g(2). \quad (40)$$

More generally the number of ancestors $m_g(k)$ common to k individuals at generation g

$$m_g(k) = \sum_{p=1}^k \frac{k!}{p! (k-p)!} (-1)^{p+1} n_g(p). \quad (41)$$

For N large and $k \ll N$, one can expand the solution of (39) in powers of $1/N$ to get

$$n_g(k) = k 2^g - \frac{k^2}{N} (2^{2g} - 2^g) + \frac{k^3}{9N^2} (8 \times 2^{3g} - 18 \times 2^{2g} + 10 \times 2^g) + \dots \quad (42)$$

Using this expansion in (41), one gets that, as long as $k 2^g \ll N$,

$$m_g(k) \sim \frac{2^{kg}}{N^{k-1}}. \quad (43)$$

By requiring that $m_g(p) \sim 1$, one then obtains the number g_k of generations needed to find at least one common ancestor to k individuals

$$g_k \simeq \frac{k-1}{k} \frac{\ln N}{\ln 2}. \quad (44)$$

This expression agrees well with simulations performed by Stephane Munier [47]. As the size k increases, it converges to Chang's expression (30).

For $2^g \sim N$, the expansion (42) can be rewritten to leading order as a scaling function

$$n_g(k) \simeq NH\left(\frac{k 2^g}{N}\right) \quad \text{with} \quad H(x) = x - x^2 + \frac{8}{9}x^3 - \frac{46}{63}x^4 + \dots \quad (45)$$

where $H(x)$ is the solution of

$$H(2x) = F[H(x)] \quad (46)$$

which starts at $x = 0$ as $H(x) = x$ and where F is given in (38). For large x , one can see from (46) that $H(x) \rightarrow x^*$, the attractive fixed point of the map $x \rightarrow F(x)$, and linearizing the map (38) around this fixed point one gets that for large x

$$H(x) \simeq x^* - B\left(\frac{\ln x}{\ln 2}\right) x^{-\alpha} \quad \text{where} \quad \alpha = -\frac{\ln(2(1-x^*))}{\ln 2} \simeq 1.3, \quad (47)$$

and where B is a periodic function $B(x) = B(x+1)$ of period 1.

To find the number of generations \widetilde{g}_k for a group of k individuals to have all their ancestors in common, one should write that $n_g(k) - m_g(k) = 0$. Here, as we use deterministic equations, the difference $n_g(k) - m_g(k)$ tends exponentially to zero without ever reaching it, so we replace this condition by $n_g(k) - m_g(k) \sim 1$ arguing that the stochastic nature of the evolution of $n_g(k)$ and $m_g(k)$ should make this difference vanish quickly after it has become of order 1. From (45,47),

$$n_g(k) \simeq NH\left(\frac{k}{N}\right) \simeq N\left[x^* - \frac{N^\alpha}{2^{g\alpha}} \frac{1}{k^\alpha} B\left(\frac{\ln k - \ln N}{\ln 2}\right)\right], \quad (48)$$

and one then gets (41) for $m_g(k)$

$$m_g(k) \simeq N\left[x^* - \frac{N^\alpha}{2^{g\alpha}} \sum_{p=1}^k \frac{k!}{p!(p-k)!} (-1)^{p+1} \frac{1}{p^\alpha} B\left(\frac{\ln p - \ln N}{\ln 2}\right)\right]. \quad (49)$$

Then, as the function B is periodic, the coefficient to N^α remains bounded in (48,49) as N becomes large and one gets that $n_g(k) - m_g(k)$ is $N^{\alpha+1}/2^{g\alpha}$ times a number of order 1. Therefore, one gets for \widetilde{g}_k :

$$\widetilde{g}_k \simeq \frac{(1+\alpha)\ln N}{\alpha \ln 2} = \left[\frac{1}{\ln 2} - \frac{1}{\ln(2(1-x^*))}\right] \ln N. \quad (50)$$

We see that for large N , the times \widetilde{g}_k do not depend on k to leading order in N and that the expression agrees with Chang's prediction (31).

5.2. An attempt to include selection

One can imagine various ways of introducing selection in a model with sexual reproduction. In an attempt to do so, we consider the same model as in the neutral case: a population of fixed size N , each individual having its two parents chosen at random in the previous generation. Then we introduce selection by saying that each parent i is chosen with a probability p_i . We assume that these p_i are all of order $1/N$, that their distribution remains the same at all generations, and that there is no correlation between the p_i of an individual and the p_i 's of its parents (the selective advantage is not inheritable). All the analysis of the neutral case can be extended to this case, the only change being that the function $F(x)$ defined in (38) becomes

$$F = \frac{1}{N} \sum_{i=1}^N \left(1 - e^{-2Np_i x}\right) \quad \text{with} \quad \sum_i p_i = 1. \quad (51)$$

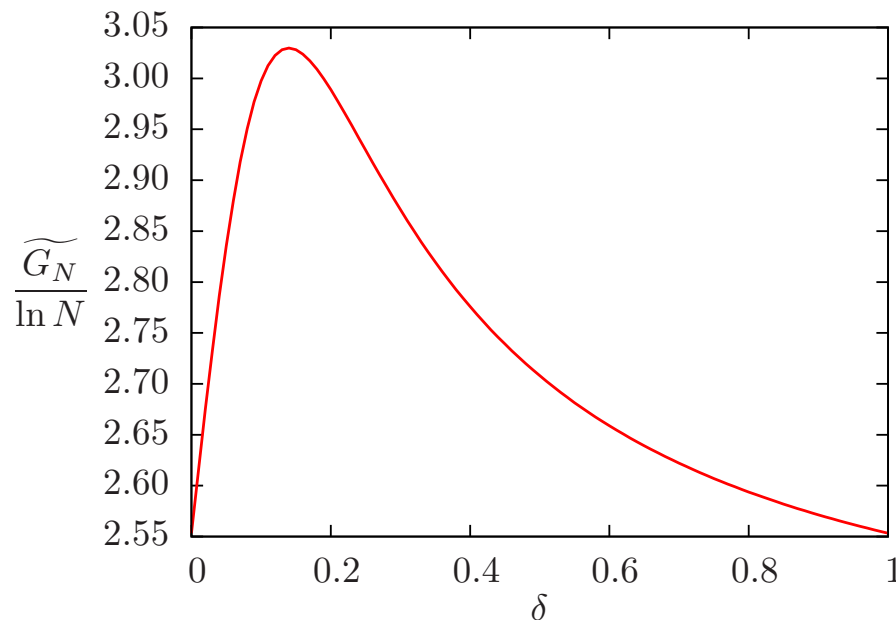


Figure 8. The time \widetilde{G}_N for the whole population to have all their ancestors in common versus δ when the p_i take two values: $1/(1+4\delta)$ with probability $1-\delta$ and $5/(1+4\delta)$ with probability δ . We see that the effect of selection (here having two possible values of the p_i is to increase the time \widetilde{G}_N while in the limit $\delta = 0$ or $\delta = 1$ one recovers Chang's value (31).

As the function F is modified, all the properties which depend on the precise shape of F , such as the fixed point x^* , are modified. In particular (31,50) become

$$\widetilde{g}_k \simeq \widetilde{G}_N \simeq \left[\frac{1}{\ln 2} - \frac{1}{\ln F'(x^*)} \right] \ln N, \quad (52)$$

while the times g_k to find at least one common ancestor remain unchanged (44).

For a particular choice of the p_i where $p_i = 5/(1+4\delta)/N$ for $i < N\delta$ and $p_i = 1/(1+4\delta)/N$ for $i > N\delta$ we show the δ dependence of \widetilde{G}_N on figure 8. We see that the time for the whole population to have all their ancestors in common is modified by selection and that, at least for the particular distribution considered here, selection has the effect of increasing the time $\widetilde{g}_k \simeq \widetilde{G}_N$. This is somewhat surprising as it goes in opposite direction of what we saw in the case of asexual evolution where the effect of selection was to shorten the coalescence times.

6. Conclusion

Here we have reviewed a series of recent results on simple models of evolution with and without selection.

In the case of asexual evolution, at least for the models we considered here in sections 3 and 4, the effect of selection is to make the coalescence times T_p (=the age of the most recent common ancestor to p individuals) grow logarithmically (24,29) with the size of the population in contrast to the linear growth of the (neutral) Wright-Fisher

model (section 2.2). Moreover the statistics of the genealogies are modified by selection and seem to be always given by the Bolthausen-Sznitman coalescent (8).

Apart for the exponential model, for which a full mathematical treatment is possible (see section 3 and [6, 7, 10]), a detailed theory, showing that the generic models of section 4 lead also to the Bolthausen-Sznitman coalescent, is still missing. Recent mathematical works, however, where selection is replaced by an absorbing wall moving at a constant velocity [48, 49, 50] or with a more complicated dynamics [51] to keep the size of the population bounded have shown that the Bolthausen-Sznitman coalescent does give the tree statistics one should see for this whole class of models.

When looking at the N values $x_i(g)$ as a cloud of points moving along the fitness axis, all the models in section 4 lead to a motion of this cloud of points very related to the motion of travelling waves of the noisy Fisher-KPP equation. In fact the same phenomenological theory [45, 7] which allows one to understand the fluctuations of the position of these travelling waves can be used for the models of section 4 and leads to the Bolthausen-Sznitman statistics of the genealogies of these models.

Other models of evolution with selection on a smooth fitness landscape have been studied in the past [41, 42, 43]. In these models each individual has an average number of offspring proportional to the exponential of its fitness (very much like in the exponential model of section 3) giving a huge advantage to the leaders. Obviously, an interesting open question would be to investigate their genealogies in order to see whether these other models lead to new statistics or if they belong to the Bolthausen-Sznitman universality class. Whether other ways of introducing selection (rugged landscapes [52], competition between a few alleles [53]) could lead to a few number of other universality classes is also an interesting question to investigate.

In the sexual case, we have recalled the main results due to Chang [1], on the numbers of generations needed to find at least one common ancestor or all common ancestors to the whole population in the neutral case. In an attempt to include selection, we have seen how these times are modified, with the rather counter-intuitive result that selection might increase the number of generations to have all ancestors common to the whole population (in contrast to what happens in the asexual case where this time decreases due to selection). Of course a more exhaustive study of the effect of selection on the genealogies, in the case of sexual reproduction, would be suitable in particular when one includes recombination [54, 55, 56].

Acknowledgments

We wish to thank S. Munier and A.H. Mueller with whom several results reviewed here were obtained.

References

- [1] J. Chang, Recent common ancestors of all present-day individuals, *Adv. in Appl. Probab.* **31**, 1002 (1999)

- [2] J.F.C. Kingman, The coalescent, *Stoch. Proc. Appl.* **13**, 235 (1982)
- [3] J.F.C. Kingman, On the genealogy of large populations, *J. Appl. Probab.* **19**, 27 (1982)
- [4] R.R. Hudson, Gene genealogies and the coalescent process, *Oxford Surveys in Evolutionary Biology* **7**, 1 (1991)
- [5] P. Donnelly, S. Tavaré, Coalescents and genealogical structure under neutrality, *Annual Review of Genetics* **29**, 401 (1995)
- [6] É. Brunet, B. Derrida, A.H. Mueller, S. Munier, Noisy traveling waves: effect of selection on genealogies, *Europhys. Lett.* **76**, 1 (2006)
- [7] É. Brunet, B. Derrida, A.H. Mueller, S. Munier, Effect of selection on ancestry: An exactly soluble case and its phenomenological generalization, *Phys. Rev. E* **76**, 041104 (2007)
- [8] E. Bolthausen, A.S. Sznitman, On Ruelle's probability cascades and an abstract cavity method, *Comm. Math. Phys.* **197**, 247 (1998)
- [9] B. Derrida, S. Manrubia, D. Zanette, On the genealogy of a population of biparental individuals, *J. Theor. Biol.* **203**, 303 (2000)
- [10] É. Brunet, B. Derrida, How genealogies are affected by the speed of evolution, *Phil. Mag.* **92**, 255 (2012)
- [11] M. Nordborg, Coalescent theory, in *Handbook of Statistical Genetics*, ed. by D.J. Balding, M. Bishop, C. Cannings (Wiley, Chichester, UK, 2001), pp. 179–212
- [12] N. Berestycki, Recent progress in coalescent theory, *Ensaio Matemáticos* **16**, 1 (2009)
- [13] J. Pitman, Coalescents with multiple collisions, *Ann. Probab.* **27**, 1870 (1999)
- [14] R. Durrett, J. Schweinsberg, A coalescent model for the effect of advantageous mutations on the genealogy of a population, *Stoch. Proc. Appl.* **115**, 1628 (2005)
- [15] T. Huillet, M. Möhle, On the extended Moran model and its relation to coalescents with multiple collisions, *Theor. Pop. Biol.* (2012), in press
- [16] S. Tavaré, Line-of-descent and genealogical processes, and their applications in population-genetics models, *Theor. Pop. Biol.* **26**, 119 (1984)
- [17] D. Simon, B. Derrida, Evolution of the most recent common ancestor of a population with no selection, *J. Stat. Mech.*, P05002 (2006)
- [18] D. Sherrington, S. Kirkpatrick, Solvable model of a spin-glass, *Phys. Rev. Lett.* **35**, 1792 (1975)
- [19] S. Kirkpatrick, D. Sherrington, Infinite-ranged models of spin-glasses, *Phys. Rev. B* **17**, 4384 (1978)
- [20] M. Mézard, G. Parisi, N. Sourlas, G. Toulouse, M.A. Virasoro, Replica symmetry breaking and the nature of the spin-glass phase, *Journal de Physique* **45**, 843 (1984)
- [21] M. Mézard, G. Parisi, M.A. Virasoro, *Spin glass theory and beyond*, vol. 9 (World Scientific Lecture Notes in Physics, 1987)
- [22] G. Parisi, Infinite number of order parameters for spin-glasses, *Phys. Rev. Lett.* **43**, 1754 (1979)
- [23] G. Parisi, The order parameter for spin glasses: A function on the interval 0-1, *J. Phys. A* **13**, 1101 (1980)
- [24] G. Parisi, Order parameter for spin-glasses, *Phys. Rev. Lett.* **50**, 1946 (1983)
- [25] R.C. Griffiths, Lines of descent in the diffusion-approximation of neutral Wright-Fisher models, *Theor. Pop. Biol.* **17**, 37 (1980)
- [26] J.M. Morel, B. Teissier, A. Etheridge, *Some Mathematical Models from Population Genetics*. Lecture Notes in Mathematics (2011)
- [27] M. Möhle, S. Sagitov, A classification of coalescent processes for haploid exchangeable population models, *Ann. Probab.* **29**, 1547 (2001)
- [28] W.J. Ewens, The sampling theory of selectively neutral alleles, *Theor. Pop. Biol.* **3**, 87 (1972)
- [29] F. Tajima, Evolutionary relationship of DNA sequences in finite populations, *Genetics* **105**, 437 (1983)
- [30] Y.X. Fu, W.H. Li, Estimating the age of the common ancestor of a sample of DNA sequences, *Mol. Biol. and Evol.* **14**, 195 (1997)
- [31] J. Degnan, L. Salter, Gene tree distributions under the coalescent process, *Evolution* **59**, 24 (2005)
- [32] B. Derrida, L. Peliti, Evolution in a flat fitness landscape, *Bull. Math. Biol.* **53**, 355 (1991)

- [33] N.A. Rosenberg, M. Nordborg, Genealogical trees, coalescent theory and the analysis of genetic polymorphisms, *Nature Rev. Genetics* **3**, 380 (2002)
- [34] L. Prignano, M. Serva, Genealogical trees from genetic distances, *Eur. Phys. J. B* **69**, 20 (2009)
- [35] M. Serva, On the genealogy of populations: trees, branches and offspring, *J. Stat. Mech.*, P07011 (2005)
- [36] B. Derrida, From random walks to spin glasses, *Physica D* **107**, 186 (1997)
- [37] J.T. Cox, Coalescing random walks and voter model consensus times on the torus in \mathbb{Z}^d , *Ann. Probab.* **17**, 1333 (1989)
- [38] V. Limic, A. Sturm, The spatial Λ -coalescent, *Elect. Journ. Prob.* **11**, 363 (2006)
- [39] E. Brunet, B. Derrida, D. Simon, Universal tree structures in directed polymers and models of evolving populations, *Phys. Rev. E* **78**, 061102 (2008)
- [40] P. Bak, K. Sneppen, Punctuated equilibrium and criticality in a simple model of evolution, *Phys. Rev. Lett.* **71**, 4083 (1993)
- [41] D.A. Kessler, H. Levine, D. Ridgway, L. Tsimring, Evolution on a smooth landscape, *J. Stat. Phys.* **87**, 519 (1997)
- [42] M.M. Desai, D.S. Fisher, A.W. Murray, The speed of evolution and maintenance of variation in asexual populations, *Current Biol.* **17**, 385 (2007)
- [43] I.M. Rouzine, E. Brunet, C.O. Wilke, The traveling wave approach to asexual evolution: Muller's ratchet and speed of adaptation, *Theor. Pop. Biol.* **73**, 24 (2008)
- [44] F. Yu, A. Etheridge, Rate of adaptation of large populations, in *Evolutionary Biology from Concept to Application*, ed. by P. Pontarotti (Springer-Verlag, Berlin Heidelberg, 2008), pp. 3–27
- [45] É. Brunet, B. Derrida, A.H. Mueller, S. Munier, A phenomenological theory giving the full statistics of the position of fluctuating pulled fronts, *Phys. Rev. E* **73**, 056126 (2006)
- [46] D. Rohde, S. Olson, J. Chang, Modelling the recent common ancestry of all living humans, *Nature* **431**, 562 (2004)
- [47] S. Munier. Private communication
- [48] H. Kesten, Branching brownian motion with absorption, *Stoch. Proc. and Appl.* **7**, 9 (1978)
- [49] B. Derrida, D. Simon, The survival probability of a branching random walk in presence of an absorbing wall, *Eur. Phys. Lett.* **78**, 60006 (2007)
- [50] J. Berestycki, N. Berestycki, J. Schweinsberg. The genealogy of branching brownian motion with absorption. *arXiv:1001.2337* (2010)
- [51] P. Maillard. Branching brownian motion with selection of the N right-most particles: an approximate model. *arXiv:1112.0266* (2011)
- [52] K. Jain, J. Krug, Evolutionary trajectories in rugged fitness landscapes, *J. Stat. Mech.*, P04008 (2005)
- [53] C. Neuhauser, S. Krone, The genealogy of samples in models with selection, *Genetics* **145**, 519 (1997)
- [54] N.H. Barton, A.M. Etheridge, The effect of selection on genealogies, *Genetics* **166**, 1115 (2004)
- [55] R.A. Neher, B.I. Shraiman, D.S. Fisher, Rate of adaptation in large sexual populations, *Genetics* **184**, 467 (2010)
- [56] R.R. Hudson, N.L. Kaplan, The coalescent process in models with selection and recombination, *Genetics* **120**, 831 (1988)